No Golden Path – A Cautionary Tale of Quality and Bias

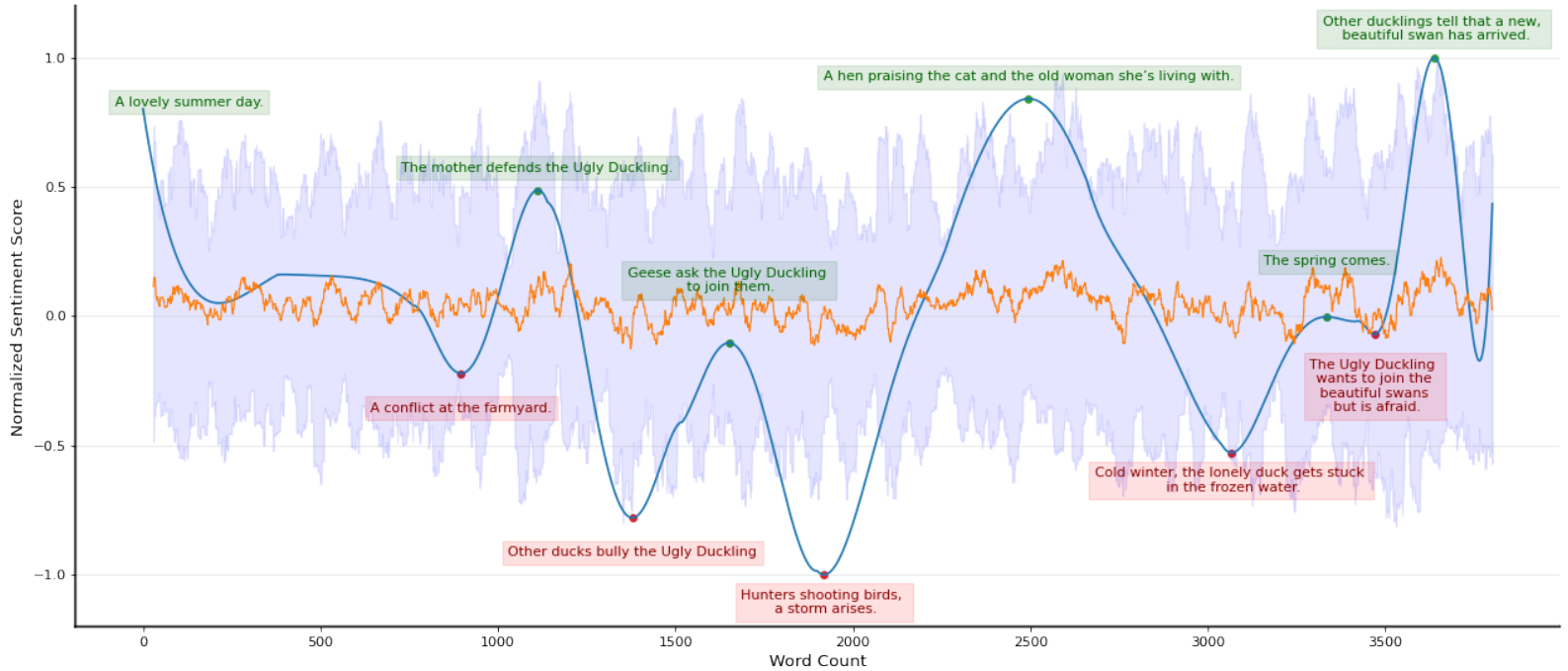
DARIAH Annual Event 2022: Storytelling

Ida Marie S. Lassen, Yuri Bizzoni, Telma Peura, Mads Rosendhal Thomsen,
Kristoffer L. Nielbo

Center for Humanities Computing AArhus|chcaa.au.dk
aarhus university, denmark

CENTER FOR HUMANITIES
COMPUTING AARHUS

Quality Assessment of Narratives

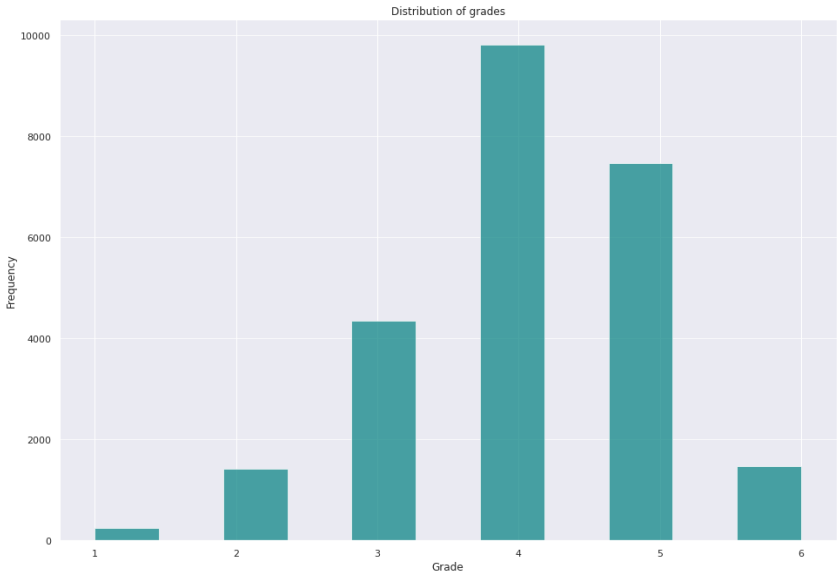


Narrative organization of information ties together storytelling in its many modalities. One archetypal expression of narratives is literary fiction.

A good ('successful') narrative has to balance the reader's motivations, that is, a story should not be too predictable (ex. 'boring') or too unpredictable (ex. 'intransparent') in order to motivate the reader to experience it.

- Model narrative, and by extension storytelling, on literary fiction.
- Explore two measures of success that are (or can be) characterized by known and unknown biases.

Extrinsic Success: Review Ratings

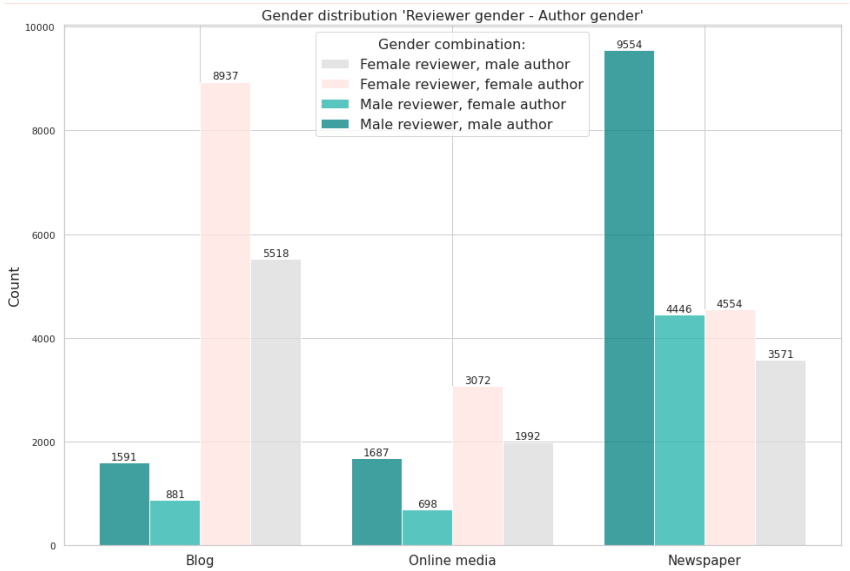


Reader reviews, a noisy but ecologically valid measure. A narrative is successful if readers rate it high. Such an 'extrinsic' success criterion is tempting because it is relatively easy to access, reflects readers' preferences in a natural setting, and its standardization appears trivial.

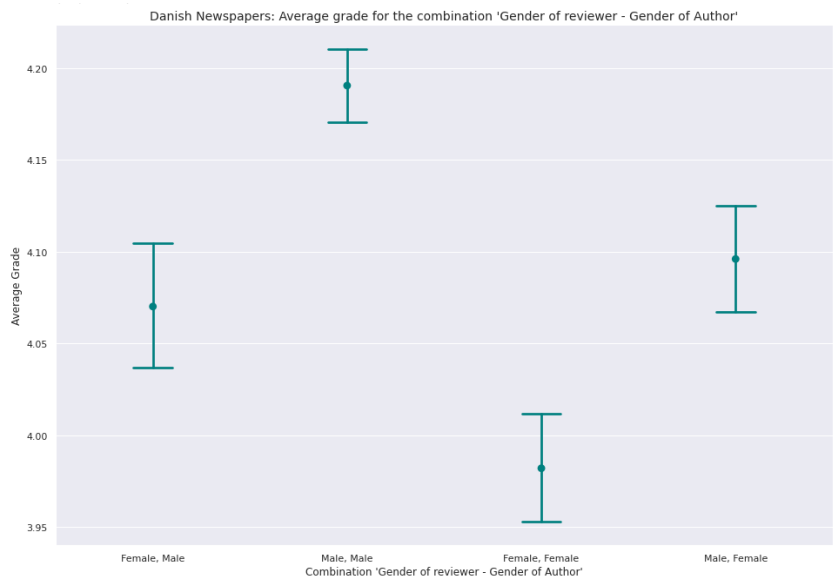
Dataset overview	
Nr of reviews	57369
Male reviewer	18958
Female reviewer	28984
Unknown	9427
Nr of different titles	14647
Male author	8056
Female author	6591
Nr of reviews by media type	
Newspapers	22131
Blogs	16791
Online media	10635
Blog-like websites	3456
Regional newspapers	2622
Weekly magazines	1566
Professional magazines	168

An overview of the dataset from bog.nu. The category Online media includes (literary) sites that fall between online newspapers and personal blogs.

Known Biases in Extrinsic Success



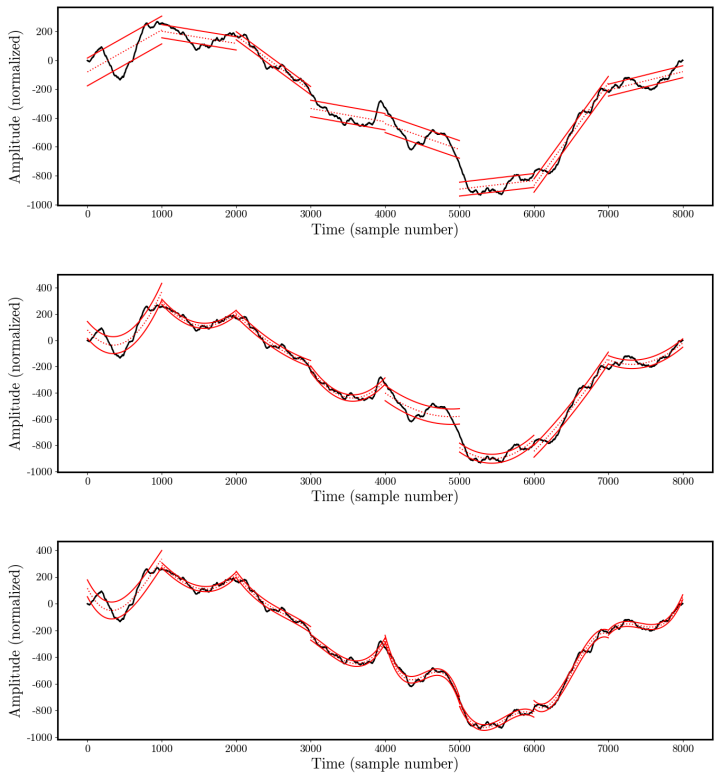
Caption



Caption

A criterion that relies on reviews is prone to several well-known biases, for instance, grading disparities between gender, ethnicity and race, which point to fairness challenges in classification of real-world data.

ADAPTIVE FILTERING



fitting of local polynomial functions for smoothing

– **partition a time series** into segments (or windows) of length $w = 2n + 1$ points, where neighboring segments overlap by $n + 1$

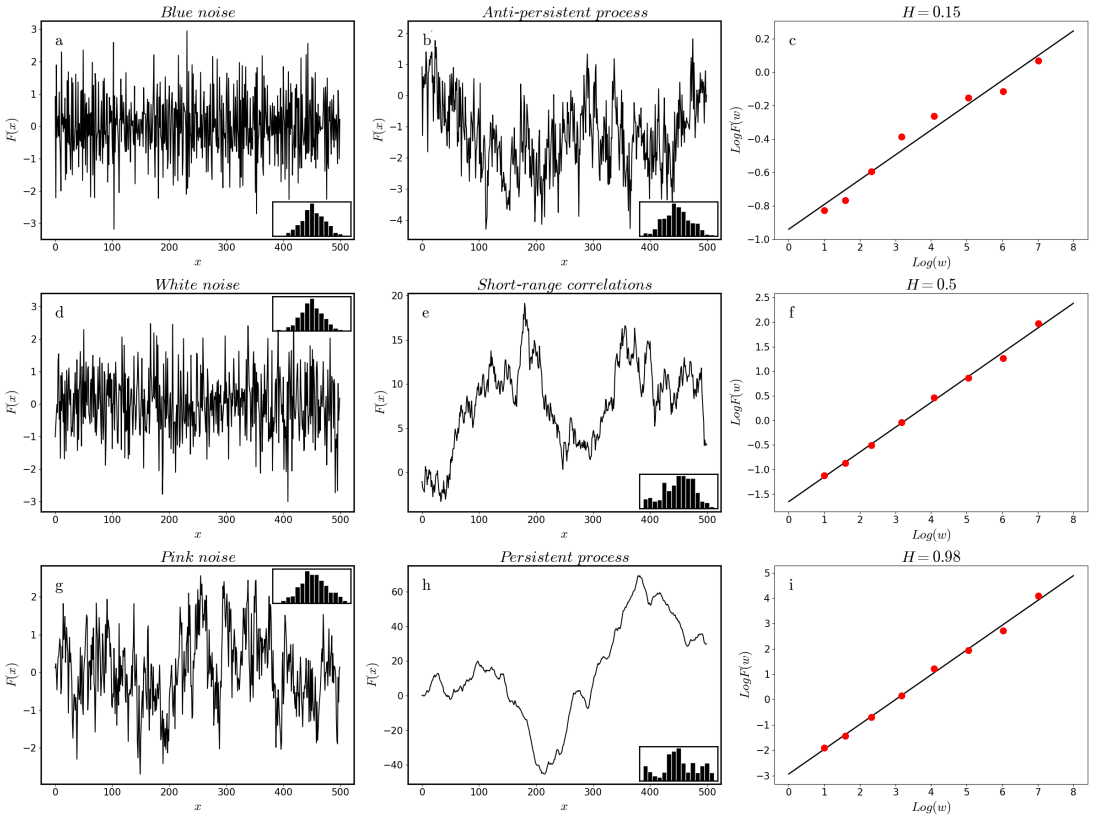
– **fit a best polynomial** of order D w. standard least-squares

– polynomials in **overlapping regions** are combined using:

$$y^{(c)}(l_1) = w_1 y^{(i)}(l + n) + w_2 y^{(i)}(l),$$
$$l = 1, 2, ..., n + 1$$

such that the **global fit** will be the best (smoothest) fit of the overall time series

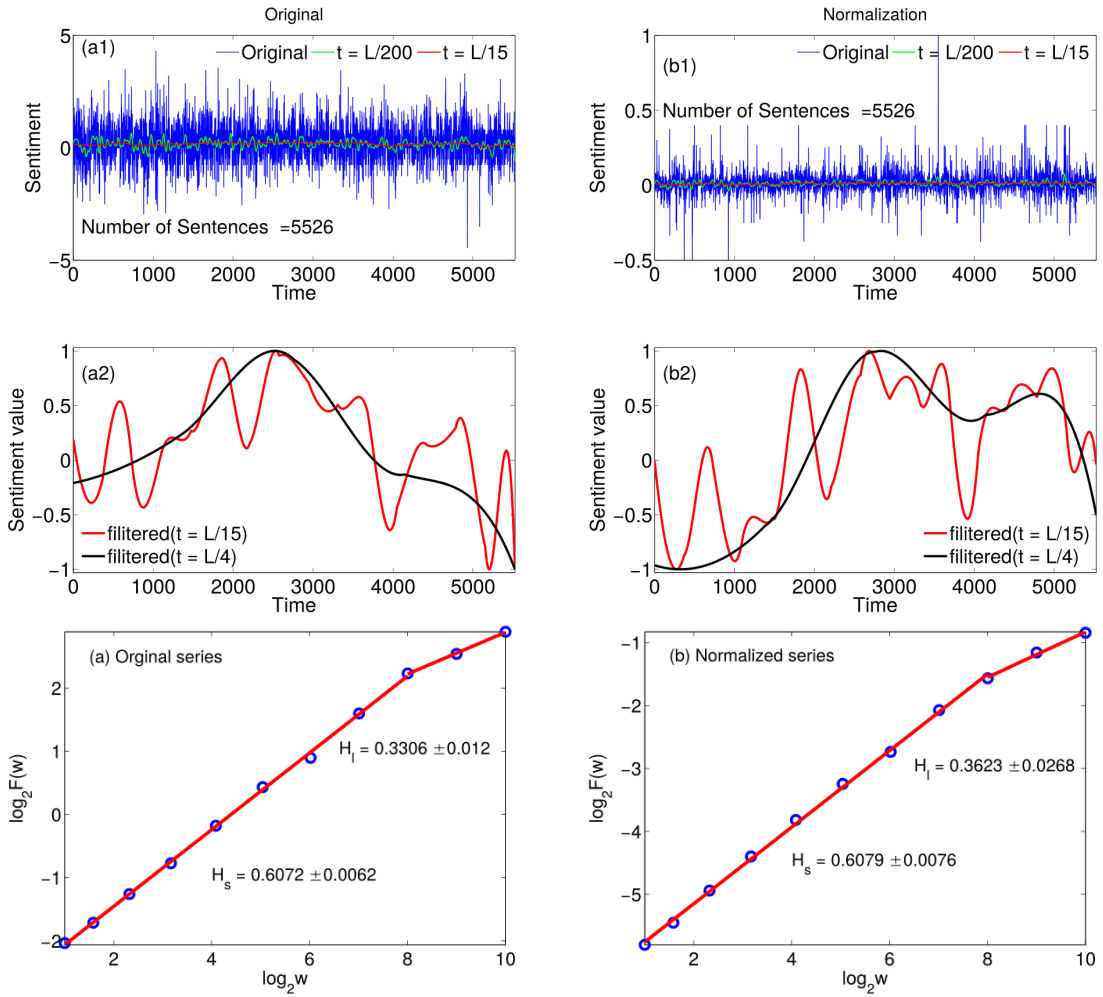
FRACTAL ANALYSIS



– **construct a random walk** $u(n) = \sum_{k=1}^{n}(x_k - \overline{x}), \quad n = 1, 2, \cdots, N,$
– divide the random walk process into **non-overlapping segments**
– determine the **local trends** of each segment as the best polynomial fit
– determine the average variance over all the segments and residual $u(i) - v(i)$ of the fit is fluctuations around global trend and its variance is the **Hurst parameter** (H)

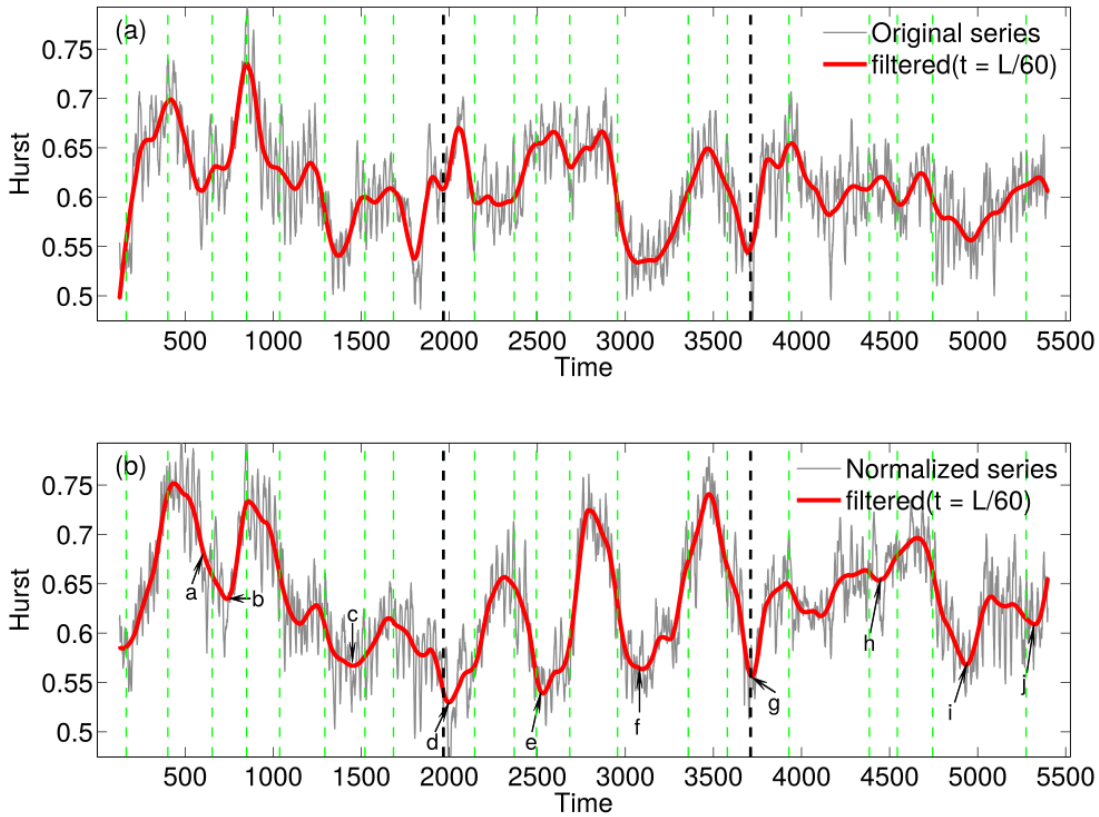
$\Rightarrow$ H **quantifies persistence** in time series: $0 < H < 0.5$ is an anti-persistent process, $H = 0.5$ is a short-memory process, and $0.5 < H < 1$ is a persistent process

INSTRINSIC SUCCESS: SELF-SIMILARITY OF STORY ARC



$0.5 < H < 1$ indicates a **coherent narrative**; $H = 0.5$ indicates a narrative that is **incoherent**, almost random (i.e., a collection of short stories); and $H < 0.5$ indicates a overly **rigid** and potentially bland narrative (i.e., a monotonous and predictable story).

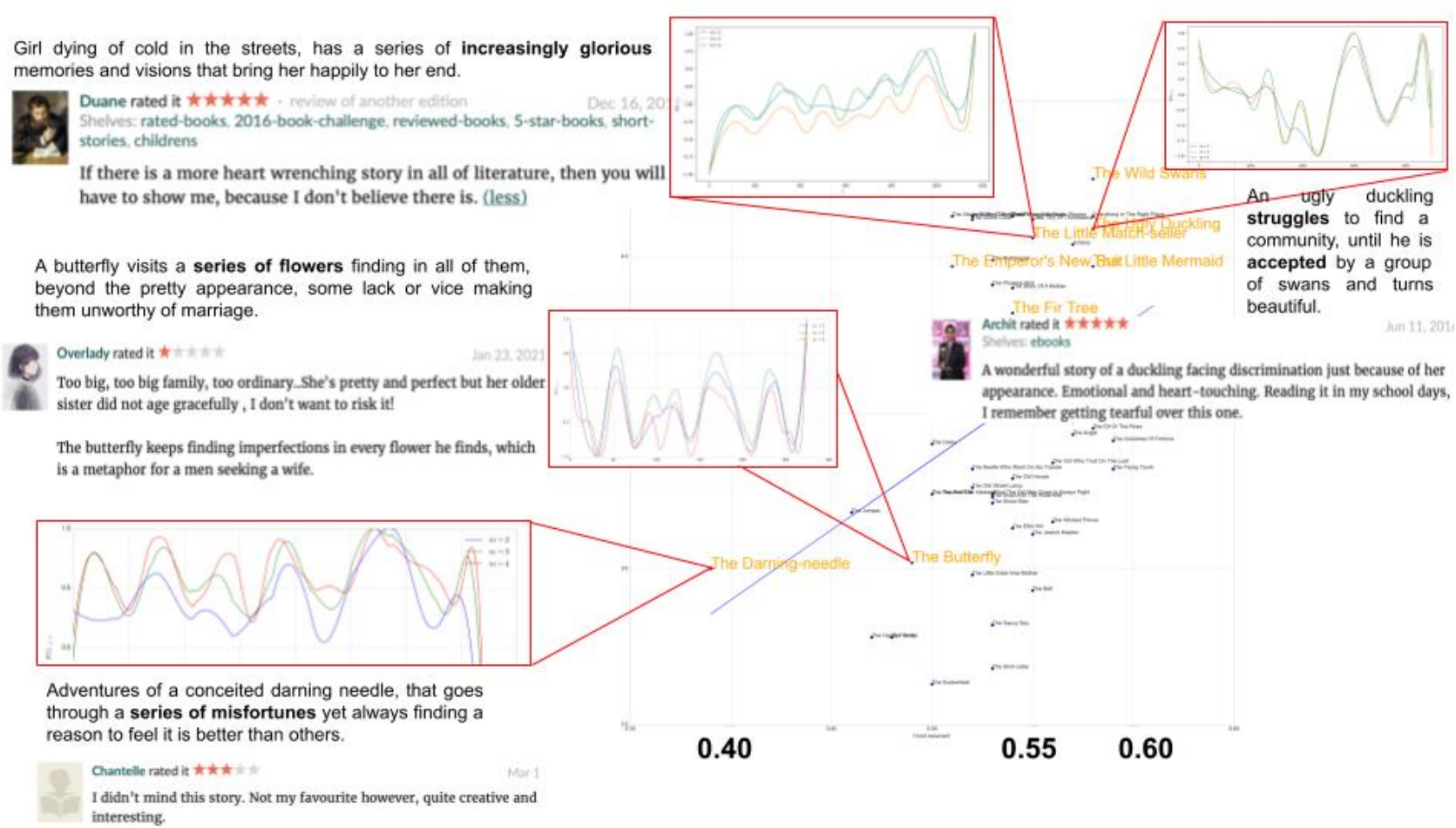
LOCAL DYNAMIC BEHAVIOR



$a - j$ indicate **change points in the narrative** (suspense), e.g., temporal shift (a: present-to-past), change in cognitive or emotional states (e.g., c: Miss Lucy informs about actual state/clones)

– **local minima** reflect disruptions or points of narrative change, positive **incline** reflect continuous (persistent) narrative development, and **decline** a movement towards disruptions

Intrinsic-Extrinsic Association



Hand-picked comments from GoodReads users are accompanied with an essential synopsis of the tale. More "zig-zag" lines reverting to the mean tend to receive less favorable overall reviews than stories having a smoother trendline. The latter ones also include many of the best known Andersen's stories. On the x-axis, the values of Hurst are evidenced: a *sweet spot* seems to lie roughly between 0.54 and 0.58.

In Conclusion

Narrative organization of information is at the core of storytelling

- A successful narrative balances the reader's motivation
- Relying on extrinsic measures of success is bias prone
- Intrinsic measures promises a context independent measure
- Intrinsic and extrinsic measures correlate
- There is *no Golden Path*, no single path that optimizes both quality assessment and bias response

→ Instead of relying on single dimensions of success, either compelling computational approaches or accessible standardization, we suggest a deliberate combination of dimensions and approaches which includes choices about bias acceptance.

```python
if questions:
    try:
        answer()
    except RunTimeError:
        pass
else:
    print('THANKS')
```

THANKS

kln@cas.au.dk
knielbo.github.io
chcaa.io

SLIDES

knielbo.github.io/files/kln_<url>.pdf

ACKNOWLEDGEMENTS

“Fabula-Net”-project funded by The Velux Foundation, NeiC’s Nordic Digital Humanities Laboratory project, and DeiC Type-1 HPC with project DeiC-AU1-L-000001. K.L. Nielbo’s work was supported by “HOPE - How Democracies Cope with COVID-19”-project funded by The Carlsberg Foundation with grant CF20-0044. The authors would like to thank bog.nu, Goodreads, and Project Gutenberg for access to data.